



Calhoun: The NPS Institutional Archive

Faculty and Researcher Publications

Faculty and Researcher Publications

2001-12-01

Convergence of Pseudospectral Discretization of Optimal Control Problems, IEEE (40th; 2001; Orlando, Florida)

Ross, Michael I.

IEEE



Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

**Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943**

<http://www.nps.edu/library>

Convergence of Pseudospectral Discretizations of Optimal Control Problems

I. Michael Ross¹
Fariba Fahroo²

Abstract

A generic nonlinear optimal control problem with a Bolza cost functional is discretized by a Legendre pseudospectral method. According to the covector mapping theorem, the Karush-Kuhn-Tucker multipliers of the discrete problem map linearly to the spectrally discretized covectors of the Bolza problem. Using this result, it is shown that the nonlinear programming problem converges to the continuous Bolza problem at a spectral rate assuming regularity of appropriate functions.

1 Introduction

In Ref.[1], Elnagar et al demonstrated that the Legendre pseudospectral method could be effectively used to discretize a Bolza problem to a nonlinear programming (NLP) problem. Further, they showed that quite accurate results were possible for low orders of discretization. Inspired by this work, we presented the symmetric covector mapping theorem in [2] and extended it for state-constrained optimal control problems in [3]. According to this theorem, the Karush-Kuhn-Tucker (KKT) multipliers map linearly to the spectrally discretized covectors associated with the optimality conditions of the Bolza problem. This suggests that the NLP indeed converges to the original Bolza problem. More precisely, we prove that the NLP converges to the optimal control problem at a spectral rate. *Spectral convergence* means that the rate of convergence for the N th order approximation is $O(\alpha^N)$, $0 < \alpha < 1$, for analytic functions and $O(N^{-m})$ for every m for C^∞ functions [4]. This is in sharp contrast to convergence problems typically encountered in traditional collocation methods like the Hermite-Simpson [5] and Runge Kutta [6] methods. In these methods, when the KKT multipliers are viewed as approximations to the costates, they are typically less accurate than the approximated states and controls although the accuracy can be maintained

for a class of Runge-Kutta methods [6]. In the Legendre pseudospectral method, we are limited only by the smoothness imposed by the problem formulation.

2 Primal-Dual Problem Formulations

Problem B: Determine the trajectory-control pair, $[-1, 1] \ni \tau \mapsto \{x \in \mathbb{R}^n, u \in \mathbb{R}^m\}$ that minimizes the Bolza cost functional,

$$J[x(\cdot), u(\cdot)] = M(x(1)) + \int_{-1}^1 F(x(t), u(t)) dt \quad (1)$$

$$\dot{x}(t) = f(x(t), u(t)), \quad (2)$$

$$\psi(x(-1), x(1)) = 0 \quad (3)$$

It is assumed the functions $M : \mathbb{R}^n \rightarrow \mathbb{R}$, $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, $\psi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^p$, are continuously differentiable with respect to their arguments.

Problem B^λ: Determine the triple, $[-1, 1] \ni \tau \mapsto \{x \in \mathbb{R}^n, u \in \mathbb{R}^m, \lambda \in \mathbb{R}^n\}$ that satisfies

$$\dot{\lambda} = -\frac{\partial \mathcal{H}}{\partial x} \quad (4)$$

$$\lambda(-1) = -\left(\frac{\partial \psi}{\partial x(-1)}\right)^T \nu \quad (5)$$

$$\lambda(1) = \frac{\partial M}{\partial x(+1)} + \left(\frac{\partial \psi}{\partial x(+1)}\right)^T \nu \quad (6)$$

in addition to Eqs.(2)-(3). Here, \mathcal{H} is the Hamiltonian defined in the usual manner, $\mathcal{H}(x, \lambda, u) = \lambda^T f + F$ and ν is the multiplier associated with the end point condition, Eq.(3). Note that we also assume that the controls are implicitly given by

$$\frac{\partial \mathcal{H}}{\partial u} = 0 \quad (7)$$

Hence Eqs.(4)-(6) remain unchanged for the Hamiltonian obtained by eliminating the controls in \mathcal{H} .

3 The Approximation Method

Let $L_N(t)$ be the Legendre polynomial of degree N on the interval $[-1, 1]$. In the Legendre pseudospec-

¹Associate Professor, Department of Aero/Astro, Code AA/Ro, Naval Postgraduate School, Monterey, CA 93943, E-mail: imross@nps.navy.mil.

²Associate Professor, Department of Mathematics, Code Ma/Ff, Naval Postgraduate School, Monterey, CA 93943, E-mail: fahroo@nps.navy.mil.

tral approximation of Eqs.(1)-(3), the Legendre-Gauss-Lobatto (LGL) points, $t_l, l = 0, \dots, N$ are used. These points are given by $t_0 = -1$, $t_N = 1$, and for $1 \leq l \leq N-1$, t_l are the zeros of \dot{L}_N , the derivative of the Legendre polynomial, L_N . In the first step of this method, the continuous variables are approximated by N th degree polynomials of the form

$$\mathbf{x}(t) \approx \mathbf{x}^N(t) = \sum_{l=0}^N \mathbf{x}_l \phi_l(t) \quad (8)$$

$$\mathbf{u}(t) \approx \mathbf{u}^N(t) = \sum_{l=0}^N \mathbf{u}_l \phi_l(t) \quad (9)$$

where, $\mathbf{x}_l := \mathbf{x}^N(t_l)$, $\mathbf{u}_l := \mathbf{u}^N(t_l)$, are the unknown coefficients, and for $l = 0, 1, \dots, N$

$$\phi_l(t) = \frac{1}{N(N+1)L_N(t_l)} \frac{(t^2-1)\dot{L}_N(t)}{t-t_l}$$

are the Lagrange polynomials of order N . To express the derivative $\dot{\mathbf{x}}^N(t)$ in terms of $\mathbf{x}^N(t)$ at the node points t_k , Eq. (8) is differentiated and the result is evaluated at t_k to obtain a matrix multiplication of the following form:

$$\mathbf{d}_k \equiv \dot{\mathbf{x}}^N(t_k) = \sum_{l=0}^N D_{kl} \mathbf{x}_l \quad (10)$$

where D_{kl} are entries of the $(N+1) \times (N+1)$ differentiation matrix \mathbf{D} ; see Ref. [7]. This facilitates the approximation of the state dynamics to the following algebraic equations

$$\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) - \mathbf{d}_k = \mathbf{0}, \quad k = 0, \dots, N$$

Next, the integral in Eq. (1) is discretized using the Gauss-Lobatto integration rule,

$$J^N[\mathbf{X}^N, \mathbf{U}^N] \approx M(\mathbf{x}_N) + \sum_{k=0}^N F(\mathbf{x}_k, \mathbf{u}_k) w_k \quad (11)$$

where $\mathbf{X}^N = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N]^T$, $\mathbf{U}^N = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_N]^T$ and w_k are the LGL weights. The boundary conditions readily simplify to $\psi(\mathbf{x}_0, \mathbf{x}_N) = \mathbf{0}$. Thus, Problem B is discretized by the following nonlinear programming problem:

Problem B^N: Find the $(N+1)(n+m)$ vector $[\mathbf{X}^N, \mathbf{U}^N]^T$ that minimizes

$$J^N[\mathbf{X}^N, \mathbf{U}^N] = \sum_{k=0}^N F(\mathbf{x}_k, \mathbf{u}_k) w_k + M(\mathbf{x}_N) \quad (12)$$

subject to

$$\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) - \sum_{l=0}^N D_{kl} \mathbf{x}_l = \mathbf{0}, \quad k = 0, \dots, N \quad (13)$$

$$\psi(\mathbf{x}_0, \mathbf{x}_N) = \mathbf{0} \quad (14)$$

The adjoint equations (4)-(6) can also be discretized in the same manner as the primal system. The costate is approximated by the N th degree polynomial:

$$\lambda(t) \approx \lambda^N(t) = \sum_{l=0}^N \lambda_l \phi_l(t) \quad (15)$$

where $\lambda_l := \lambda^N(t_l)$. The discretized Hamiltonian is expressed as

$$\mathcal{H}_i^N = \mathcal{H}^N(\mathbf{x}_i, \mathbf{u}_i, \lambda_i) = \lambda_i^T \mathbf{f}(\mathbf{x}_i, \mathbf{u}_i) + F(\mathbf{x}_i, \mathbf{u}_i) \quad (16)$$

Thus, discretization of Problem B^λ is given by [3]

Problem B^{λN}: Find the coefficients $\{\mathbf{x}_k\}$, $\{\mathbf{u}_k\}$, $\{\lambda_k\}$, and ν that satisfy the following nonlinear algebraic relations for $k = 0, \dots, N$:

$$\frac{\partial \mathcal{H}_k^N}{\partial \lambda_k} - \sum_{i=0}^N D_{ki} \mathbf{x}_i = \mathbf{0}, \quad (17)$$

$$\frac{\partial \mathcal{H}_k^N}{\partial \mathbf{x}_k} + \sum_{i=0}^N D_{ki} \lambda_i = \mathbf{0} \quad (18)$$

$$\frac{\partial \mathcal{H}_k^N}{\partial \mathbf{u}_k} = \mathbf{0} \quad (19)$$

$$\psi(\mathbf{x}_0, \mathbf{x}_N) = \mathbf{0} \quad (20)$$

$$\lambda_0 = - \left(\frac{\partial \psi}{\partial \mathbf{x}_0} \right)^T \nu \quad (21)$$

$$\lambda_N = \frac{\partial \mathcal{M}}{\partial \mathbf{x}_N} + \left(\frac{\partial \psi}{\partial \mathbf{x}_N} \right)^T \nu \quad (22)$$

4 The Main Theorem

In Ref. [3] it has been shown that

Theorem 4.1 Let $\{\mathbf{x}_i, \mathbf{u}_i, \nu, \lambda_i\}^N$ be a solution to Problem B^{λN}. Then, $\{\mathbf{x}_i, \mathbf{u}_i\}^N$ is a KKT point to Problem B^N with KKT multipliers $\{\nu, \lambda_i\}^N$.

This is the symmetric covector mapping theorem; in this paper, we show the convergence of the approximations. To show this, we consider the Sobolev space $H^m(-1, 1)$ which consists of the vector-valued functions on $[-1, 1]$ whose j th derivative $0 \leq j \leq m$ belongs to $L^2(-1, 1)$.

Theorem 4.2 Assume the state and costate variables belong to the space $H^m(-1, 1)$. Also, assume that the gradients of the Hamiltonians with respect to the states and costates belong to the space $H^{m-1}(-1, 1)$. Then, Problems B^N and B^{λN} approximate Problems B and B^λ to spectral accuracy so that $\{\nu, \lambda_i\}^N$ converge to the covectors $\nu, \lambda(t)$ at the rate of $O(N^{-m})$.

4.1 Proof of the Theorem

The proof of the theorem follows quite directly from standard results in spectral methods [7] and the symmetric covector mapping theorem. Essentially, we prove that the boundary value problem (BVP), Problem $B^{\lambda N}$ converges to Problem B^λ and hence from Theorem 4.1 it follows that Problem B^N converges to Problem B .

In the Legendre pseudospectral method, the interpolating polynomial $I_N y$ of a function y on $[-1, 1]$ is an element of \mathcal{P}_N , the space of N th degree polynomials. It is easily seen that the interpolant $I_N y$ is the projection of y upon \mathcal{P}_N with respect to the discrete inner-product $\langle y, v \rangle_N = \sum_{j=0}^N y(t_j) v(t_j) w_j$, where t_j 's are the LGL points and w_j 's are the LGL weights [7]. The basic idea of the convergence proof relies on the following interpolation estimates [7]:

Theorem 4.3 For $y \in H^m(-1, 1)$, and $0 \leq l \leq m$ the following estimate holds:

$$\|y - I_N y\|_{H^l(-1,1)} \leq C N^{1/2} N^{2l-m} \|y\|_{H^m(-1,1)}.$$

Here and in the remainder of this paper, we use C as generic constants (not all the same) independent of N . A particular case of the last inequality is the following estimate of the error between the exact and Legendre collocation differentiation:

$$\|y'(t) - (I_N y)'(t)\|_{L^2(-1,1)} \leq C N^{5/2-m} \|y\|_{H^m(-1,1)} \quad (23)$$

Now, let $y(t)$ be the exact solution to the BVP

$$\dot{y} = g(y), \quad B(y(-1), y(1)) = 0 \quad (24)$$

where $g: \mathbb{R}^r \rightarrow \mathbb{R}^r$ and $B: \mathbb{R}^r \times \mathbb{R}^r \rightarrow \mathbb{R}^r$. From the abstract framework presented in Ref. [7] (Sec. 10.4), one has the following convergence estimate for the discrete pseudospectral solutions, y^N :

$$\frac{\|y - y^N\|_E + \langle g(y), e \rangle - \langle I_N g, e \rangle_N}{\|e\|_E} \leq C \|y - I_N y\|_E + \quad (25)$$

where, $e = y^N - I_N y$ and E is the state space. For our problem, $y = [x, \lambda]^T$ and $g = [\frac{\partial \mathcal{H}}{\partial \lambda}, -\frac{\partial \mathcal{H}}{\partial x}]^T$. From the uniform equivalence of the discrete and the continuous L^2 norms, all of the error estimates described above hold if the L^2 norms of the errors are replaced by the discrete ones at the interpolation points. Thus, from Theorem 4.3, it is apparent that we have the following result for $E = H^m(-1, 1)$:

$$\|y - y^N\|_{H^l(-1,1)} \leq C N^{1/2} N^{2l-m} (\|y\|_{H^m(-1,1)} + \|g\|_{H^{m-1}(-1,1)}) \quad (26)$$

A particular case of this estimate is the L^2 -norm error between the pseudospectral solution and the exact solution: this error goes to zero at a spectral rate under the stated assumptions on regularity of y and g . QED.

It is useful to note that the convergence properties of pseudospectral methods are quite good even for discontinuous functions as noted in [7]. Numerical evidence (see Refs.[1]-[3]) supports this observation for optimal control problems as well.

Note that we *inherit* the convergence properties from the smoothness of the relevant functions (controls, states and costates) and the associated maps (i.e. f and $\frac{\partial \mathcal{H}}{\partial x}$). Thus, for any given problem, if one can ascertain the precise class of the functions that define the problem, the convergence of the pseudospectral discretization can be determined by this theorem.

References

- [1] J. Elnagar, M. A. Kazemi, and Razzaghi, M., "The pseudospectral Legendre method for discretizing optimal control problems," *IEEE Transactions on Automatic Control*, Vol. 40, No. 10, pp. 1793-1796, 1995.
- [2] F. Fahroo and I. M. Ross, "Costate estimation by a Legendre pseudospectral method," *Proceedings of the 1998 AIAA GNC conference*, in Boston, MA, August 1998, Paper 98-4222, also in the *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, pp. 270-277, 2001.
- [3] I. M. Ross and F. Fahroo, "A pseudospectral transformation of the covectors of optimal control systems," *Proceedings of the 1st IFAC Symposium on System, Structure and Control*, Prague, Czech Republic, August 2001.
- [4] L. N. Trefethen, *Spectral Methods in MATLAB*, SIAM, Philadelphia, PA, 2000.
- [5] P. J. Enright and B. A., Conway, "Discrete approximations to optimal trajectories using direct transcription and nonlinear programming," *Journal of Guidance, Controls and Dynamics*, Vol. 15, No. 4, pp. 994-1002, 1992.
- [6] W. W. Hager, "Runge-Kutta methods in optimal control and the transformed adjoint system," *Numerische Mathematik*, 87, pp. 247-282, 2000.
- [7] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in Fluid Dynamics*. Springer Verlag, New York, 1988.